GARCH-based Volatility Forecasts for Implied Volatility Indices

Massimiliano Cecconi Giampiero M. Gallo*
Marco J. Lombardi

Prepared for the 3rd Workshop on Mathematical Finance
Verona, Jan. 31–Feb. 1, 2002

ABSTRACT

Volatility forecasting is one of the main issues in the financial econometrics literature. Besides the many statistical models proposed throughout the years to estimate and forecast conditional variance, professional operators may rely on alternative indices of volatility supplied by banks, consultants or financial analysts. Among those indices, one of the most widely used is the so-called VXN, computed using the implied volatility of the options written on the NASDAQ–100 Index that is supplied by CBOE since 1995. In this paper we show how forecasts obtained with traditional GARCH–type models can be used to forecast the volatility index VXN.

I Introduction

The concept of volatility is somewhat elusive. In a statistical framework, it is usually associated to (the square root of) the conditional variance estimated on the basis of the information available in t, and projected τ periods ahead. On the other hand, practitioners are accustomed to seeing alternative measures, such as realized volatility and volatility measures implicit in the market prices of derivatives, such as implied volatility in option prices. This alternative approach is all the more important since there are traded financial products such as the VIX™and the VXN™which are experiencing growing success, especially among professional operators. Those indices are computed and provided on a 60–second basis as an average of implied volatilities in at–the–money options with a residual time–to–maturity equal to 30 days. The first one refers to options written on the Standard & Poor 100 Index, whilst the latter considers the NASDAQ–100.

For risk and portfolio managers, the usefulness of volatility forecasts is self-evident: it can be used to evaluate and hedge against risk, to price derivatives or to compute measures of value at risk. It is quite clear that the accuracy of the forecasts must cope with

*Corresponding author: Dipartimento di Statistica "G. Parenti", Università degli studi di Firenze – Viale G.B. Morgagni 59 – 50134 Firenze (Italy) – Tel. +39/055/4237201, FAX +39/055/4223560, E-mail: gallog@ds.unifi.it.

the quickness of computation and the easiness of availability. This is one of the reasons of the success of VIX and VXN indices which reflect in real time the market evaluation of the prevailing volatility present on stock indices.

Our study intends to investigate whether the two streams of volatility estimation can be bridged together. In particular, we want to see whether the estimation of GARCH–type models and the projection of volatility predictions thirty days ahead is of any use to forecast the volatility index VXN. With a different goal in mind, Engle and Gallo (2001) address similar issues in reference to the VIX Index.

After briefly introducing the logic of the construction of the VXN Index and the GARCH formulations (the exponentially decaying version following from Bollerslev, 1986, and the hyperbolically decaying version following Baillie, Bollerslev and Mikkelsen, 1996) that will be used, we will proceed comparing the forecasts on the volatility of the NASDAQ–100 Index obtained with various GARCH–type models with those supplied by the CBOE's VXN. The comparison will be both on the estimated volatilities and on the effects of the estimates obtained with GARCH models on the VXN series both in– and out–of sample. Concluding remarks will follow.

II GARCH Models

In this section, we intend to summarize the main features of some model specifications in the GARCH family, with specific reference to a discussion of the forecasting capabilities in the medium run (30 days). As well known, GARCH–type models capture the stylized fact observed in the data of volatility clustering and are able to reproduce the other empirical regularity of fat tails in the unconditional distributions. The most recent data available on the innovation and the previous estimate of the conditional variance is exploited in order to forecast future volatility. In this respect, the speed at which the information available at time t decays becomes essential, given the goal of attaining meaningful forecasts for 30–days ahead volatility. If the speed of convergence to the unconditional variance were to be too fast, in fact, there would be little or no gain from using conditional variance models. In what follows, we will perform a main comparison between GARCH (possibly augmented by information other than the past of the series, but still available at time t) and FIGARCH models given that the two classes of models treat the concept of persistence and long memory in a different fashion.

Given a series of asset returns r_t, conditionally modelled as

$$r_t|I_{t-1} = \mu_t + \epsilon_t,$$

where the conditional mean μ_t may contain a constant, some dummy variables to capture calendar anomalies and possibly an autoregressive or moving average term. Following the survey by Bollerslev, Engle and Nelson (1994), a GARCH class of models can be expressed as

$$\epsilon_t|I_{t-1} \sim N(0, \sigma_t^2) \tag{1}$$

where I_{t-1} denotes the information set available at time $t-1$ and σ_t^2 is a time–varying variance. Different specifications of σ_t^2 as a deterministic function of past innovations and past conditional variances give rise to several kinds of GARCH–type models: if we consider σ_t^2 as a linear function both of the p past squared innovations and of the q lagged conditional variances, we have the standard GARCH(p,q) model introduced by Bollerslev (1986)

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \equiv \omega + \alpha(L)\epsilon_t^2 + \beta(L)\sigma_t^2 \tag{2}$$

and where L denotes the lag operator, $L^i y_t \equiv y_{t-i}$. Imposing the restrictions $\beta_j = 0$ for any j, the equation (2) gives the original ARCH (p) model introduced by Engle in 1982. As noted by Bollerslev, Chou and Kroner (1992), in practice, a GARCH(1,1) specification appears to fit most financial time series well. The philosophy behind the GARCH modelling is to capture the stylized facts emerging from the empirical analysis of the series: next to volatility clustering and leptokurtosis, therefore, other suggestions have been advanced in recent years to take into account other relevant stylized facts. The different behavior of volatility to good and bad news (the so–called leverage effect) is translated into asymmetric GARCH models such as TGARCH (Glosten *et al.*, 1993 and Zakoian, 1994) where the conditional variance reacts differently to the innovations according to whether they have a positive or a negative sign:

$$\sigma_t^2 = \omega_1 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-j}^2 + \gamma S_{t-1}^- \epsilon_{t-1}^2 \tag{3}$$

where S_{t-1}^- is a dummy variable which takes value 1 if $\epsilon_{t-1} < 0$ and 0 otherwise.

Other authors (cf. Hamilton and Susmel, 1994; Gray, 1996) have advanced the idea that volatility is subject to regimes (e.g. high and low) which can be modelled as a Markov process, discretely switching between two (or more) (G)ARCH specifications. The same idea was applied independently by Hagerud (1996) and González–Rivera (1998), and developed by Anderson, Nam and Vahid (1999), in a Smooth Transition framework (Teräsvirta, 1994). Intuitively, the model considers different states of the world or regimes and allows for the possibility that the future behavior of the conditional variance depends on the current regime. As consequence, a different model will be active in different states and the regime–switching behavior is determined by a continuous function which arises to a dynamical pattern similar to that observed in the STAR models for the conditional mean equation. Formally, the *Asymmetric Nonlinear Smooth Transition* GARCH (ANST–GARCH) is

$$\sigma_t^2 = \omega_1 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + (\omega_2 + \alpha_2 \epsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2) \times F(\epsilon_{t-1}; \lambda) \tag{4}$$

where $F(\epsilon_{t-1}; \lambda)$ indicates the logistic function,

$$F(\epsilon_{t-1}; \lambda) = [1 + \exp(-\lambda \epsilon_{t-1})]^{-1}, \lambda > 0 \tag{5}$$

which is bounded between 0 and 1, and depends on the transition variable ϵ_{t-1} and on the smoothness parameter λ restricted to be strictly positive in order to avoid identification problems. Note that for $\lambda = 0$, $F(\epsilon_{t-1}; \lambda) = 1/2$ arising to a one state, or linear GARCH model and when $\lambda \to \infty$, $F(\epsilon_{t-1}; \lambda)$ becomes a step–function indicating an abrupt change between the two regimes, as in the TGARCH models[1].

[1] See Anderson *et al.* (1999) for details.

Another empirical regularity[2] emerged in recent years is the speed of decay of autocorrelations in squared innovations, which is at times badly approximated by exponential models such as GARCH. The *Fractionally Integrated* GARCH models suggested by Baillie, Bollerslev and Mikkelsen (1996) merge some of the results on fractional integration and the GARCH literature by allowing the autocorrelations to decay at a slower, hyperbolic, speed. The basic specification adopted is

$$\psi(L)(1-L)^d\epsilon_t^2 = \omega + [1-\beta(L)]w_t, \tag{6}$$

where $(1-L)^d$ denotes the *fractional difference operator*, that may be expanded in a McLaurin series to produce an infinite polynomial in L, the lag operator:

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} L^k = \sum_{k=0}^{\infty} \varphi_k(d) L^k.$$

The FIGARCH model has proven to be very useful, especially for what concerns high–frequency financial time series (see, for example, Andersen and Bollerslev, 1997). Yet, since we need to work with very long time series in order to properly capture the long–run dependency, the estimation of FIGARCH processes requires very long loops and turns out to be very time–consuming. Employing the analytic expressions for the first– and second–order derivatives presented in Lombardi and Gallo (2001) allows to save time and improve the quality of the estimates. One of the major drawbacks when working with FIGARCH processes is that its properties are not very clear yet. While consistency and stationarity have recently been proved by Caporin (2001), an interesting alternative approach to long–memory volatility processes that seems to shed some light on several open questions is provided in Davidson (2001).

Other specifications are available which will not be recalled in detail here to deal with other features of observed series which are not suitably captured by the GARCH(1,1) model (see Engle and Patton, 2001 for a recent survey).

III Empirical Features of the NASDAQ–100 Index

As mentioned above, the NASDAQ–100 Index is a weighted average of the stock prices of the top–capitalized 100 non financial companies traded at the National Association of Security Dealers Automated Quotes. The weights are updated every three months with a proprietory algorithm which is not distributed and hence there is an element of arbitrariness in the way the index is built. It is well known that the NASDAQ is a more volatile market relative to the New York Stock Exchange due to some sectors that are connected to the Information Technology evolution of recent years with a great deal of companies experiencing fast market capitalization but also more severe downturns relative to the more traditional sectors. The interest in studying such an index is twofold: on the one hand, it may characterize the behavior of the US economy in growing strategic sectors where knowledge and "immaterial capital" are very important; on the other, the index has attracted the interest of several risk managers for the wide array of derivative

[2]Cf. Ding *et al.* (1993).

products written on the index. We display in Figure 1 the NASDAQ–100 Index for the sample we used to estimate different GARCH models (Oct. 1, 1985 to Jan. 18, 2002 - a total of 4254 observations, leaving the last 67 for out–of–sample forecast comparison purposes). It is clear that the index has experienced a wild growth and increased volatility in the past couple of years: between 1998 and the beginning of 2001 the formation of the "dot com" bubble and its subsequent burst is apparent with interesting features of the volatility behavior in those years relative to the previous period. Of course, the interesting question becomes one of whether the parametric volatility models of the GARCH family are capable of capturing what may seem two different regimes in the series, at least as far as conditional second moments are concerned.

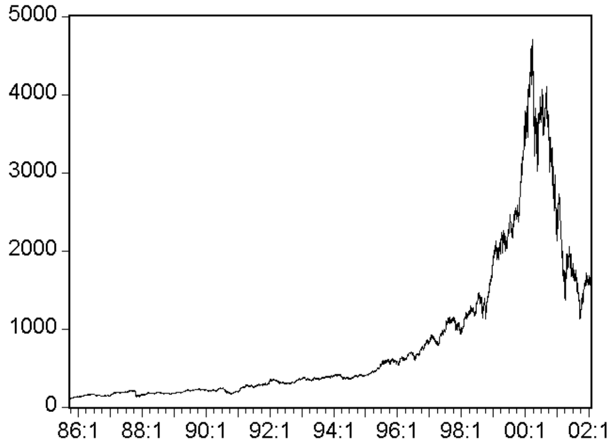

Figure 1: NASDAQ–100 Index from January 10, 1985 to January 18, 2002

In order to gain some insights on the characteristics of the series, it is instructive to look at some descriptive statistics. As customary in financial time series, the unit root hypothesis is tested with the Augmented Dickey Fuller test: here we chose four lags for the innovation process. The ADF test statistic $\hat{\tau}_\mu$ for the coefficient of the lagged level in the test equation for the first differences where a constant is present is equal to -1.08, well within the acceptance region marked by a critical value equal to -2.86 at the 5% significance level.

Working with log differences, i.e. with an approximation to daily returns, is warranted by the unit root results. What is relevant, therefore, is to characterize the modelling needs of the return series. Some descriptive statistics are reported in Table 1, where next to the characteristics of the unconditional distribution we report the test statistics for autocorrelation, ARCH and Jarque–Bera normality tests.

From the diagnostics on the returns, it is clear that some autocorrelation is present in the returns (a common occurrence for indices) and that ARCH effects need to be taken into account. Normality in the unconditional distribution is strongly rejected, but even this occurrence is not worrisome, as it is common to most daily financial time series. The modelling strategy for the conditional mean is therefore to insert an AR(1)

Mean	0.062	
Median	0.106	
Maximum	17.20	
Minimum	-16.34	
Std. Dev.	1.815	
Skewness	-0.171	
Kurtosis	11.03	
Test	Statistic	p–value
AR(12)	33.04	0.0009
ARCH(4)	538.63	0.000
Normality	11451.85	0.000

Table 1: Descriptive statistics of the NASDAQ–100 Index

term for all models,

$$r_t = \mu + \phi r_{t-1} + \epsilon_t$$

concentrating afterwards on the various ways to model conditional variance of ϵ_t according to the procedures described above.

The estimation results together with the standardized residuals diagnostics are reported in Table 2. The GARCH model exhibits the customary order of magnitude for the $\hat{\alpha}_1$ (around 0.1) and $\hat{\beta}_1$ (around 0.9) that one obtains with daily financial data. Asymmetric effects are well captured by the $\hat{\gamma}$ coefficient which, estimated in the TGARCH model, is significant and shows how conditional variance increases when past innovations are negative.

Since the estimation of ANST–GARCH is less frequent, it is advisable to spend a few words on the interpretation of the estimated coefficients. First of all the coefficient ruling the transition function, λ, is positive: its significance cannot be evaluated with a standard t-ratio statistic since its distribution is not standard under the null hypothesis (a case of nuisance parameter identified only under the alternative which creates a wide array of inferential problems, cf. Hansen, 1996). At any rate, the function can be graphed as in Figure 2 to show the sensitivity of the estimated transition to lagged innovations. The switching–regime behavior seems to be very fast, indicating that a threshold model can suitably capture the dynamic of the NASDAQ–100 Index.

The estimated coefficients are of somewhat more difficult interpretation: the coefficients of the second regime are not restricted to be all positive in order to ensure a positive and stationary conditional variance. Following Anderson *et al.* (1999), the only restrictions we impose are $\hat{\omega}_1 + 0.5\hat{\omega}_2 > 0$, $\hat{\alpha}_1 + 0.5\hat{\alpha}_2 > 0$ and $\hat{\beta}_1 + 0.5\hat{\beta}_2 > 0$ for the positivity of the conditional variance and $0 < (\hat{\alpha}_1 + 0.5\hat{\alpha}_2 + \hat{\beta}_1 + 0.5\hat{\beta}_2) < 1$ for the stationarity. All these conditions are satisfied in the estimated model.

As for the FIGARCH estimation results, the estimated order of fractional integration $\hat{d}$ is equal to 0.426, a value which is in line with other estimations on similar time series. The coefficient ψ is similar but structurally different from α_1, hence its estimated value is not directly comparable; $\hat{\beta}_1$ is lower than what is obtained in the GARCH case. This discrepancy is caused by the fact that, in the FIGARCH model,

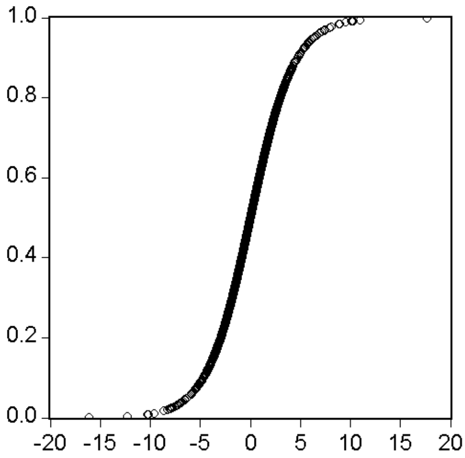

Figure 2: Estimated transition function for the ANST–GARCH model

the dependency is captured by the φ_k coefficients of the infinite lag polynomial, thus reducing the strength of the first-order dependency captured by the β_1 coefficient.

To gain further insights on the nature of the estimated conditional variances, let us compare the series with one another within sample and let us start by showing the time series profile of the conditional variance estimated with the GARCH(1,1) model in Figure 3. A meaningful comparison with the other results is achieved by showing the in–sample scatters of the estimated conditional variances for all pairs of results: this is done in Figure 4.

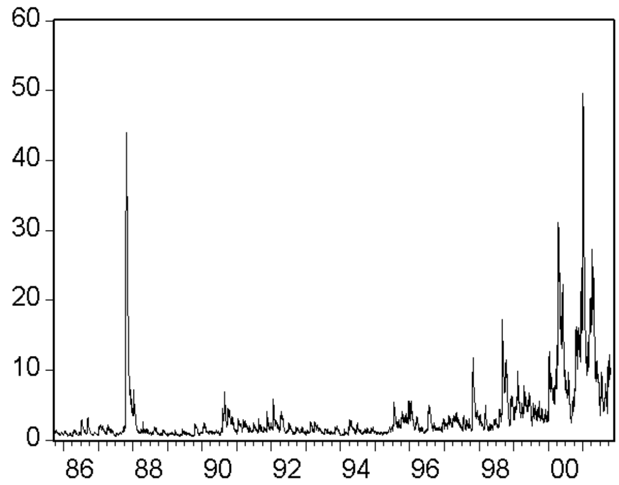

Figure 3: Conditional variance estimated in–sample – GARCH(1,1) model

	GARCH	*TGARCH*	*ANST–GARCH*	*FIGARCH*
μ	0.099	0.065	0.069	0.044
	(0.021)	(0.021)	(0.0223)	(0.020)
ϕ	0.070	0.077	0.057	0.064
	(0.016)	(0.016)	(0.013)	(0.016)
ω_1	0.033	0.040	3.550	0.091
	(0.011)	(0.011)	(0.740)	(0.034)
α_1	0.092	0.047	0.098	–
	(0.019)	(0.014)	(0.017)	
ψ	–	–	–	0.230
				(0.090)
β_1	0.897	0.895	0.452	0.528
	(0.018)	(0.014)	(0.052)	(0.103)
γ	–	0.086	–	–
		(0.028)		
ω_2	–	–	-4.850	–
			(1.470)	
α_2	–	–	0.093	–
			(0.040)	
β_2	–	–	-0.201	–
			(0.088)	
λ	–	–	0.842	–
			(0.307)	
d	–	–	–	0.426
				(0.064)
$Log-lik$	-7433.97	-7411.17	-7433.56	-7434.21
AIC	3.555	3.545	3.557	3.553
SC	3.563	3.554	3.570	3.563
Skewness	-0.442	-0.411	-0.235	-0.398
Kurtosis	5.214	4.844	6.571	4.841
Normality	991.92	711.44	2262.7	701.4
p-value	(0.000)	(0.000)	(0.000)	(0.000)
AR(12)	8.063	6.485	23.905	10.860
p-value	(0.780)	(0.889)	(0.021)	(0.541)
ARCH(4)	9.686	3.154	131.2	3.724
p-value	(0.046)	(0.532)	(0.000)	(0.445)

Table 2: Estimation results on various GARCH models

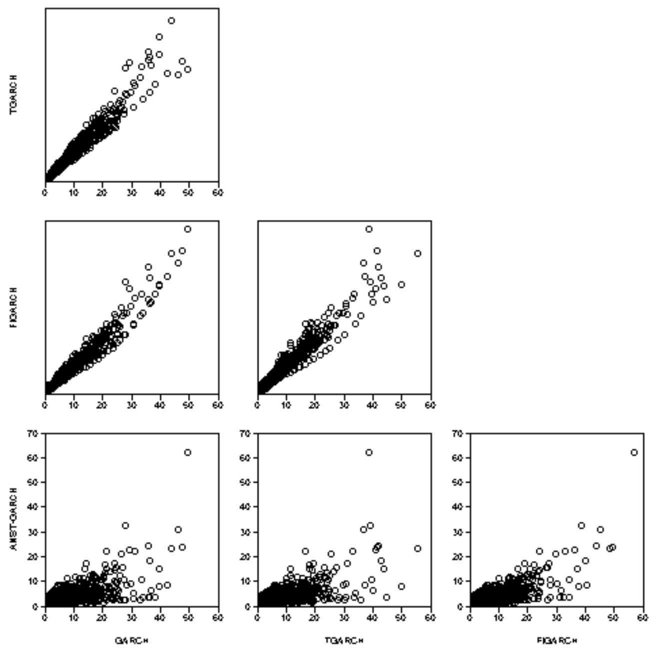

Figure 4: Scatterplot of the estimated conditional variances

At first glance, the behaviors of GARCH, TGARCH and FIGARCH (top three scattergrams) are more similar to one another than relative to the ANST–GARCH (bottom three). For the former, in fact, the scatter shows a substantial agreement in paired values which follow an increasing pattern: we do not recognize any systematic under or overestimation of the conditional variance according to either method. The case of the ANST–GARCH is somewhat different, in that it seems that the estimated values tend to be systematically lower than the ones obtained with the other models. The characteristic of this model (as in the case of the SWARCH of Hamilton and Susmel, 1994) is in a greater absorption of the shocks on the part of a change in regime (the speed of which is ruled by the coefficient $\hat{\lambda}$). The estimated persistence of the conditional variance is therefore much smaller: we can expect the estimated coefficients for the first set of models to be higher than the one for ANST–GARCH and closer to one another.

In order to evaluate this conjecture, let us consider the following regression:

$$h_t^i = c + \zeta h_{t-1}^i + u_t$$

where h_t^i is the estimated variance according to model i. Being an autoregression, one measure of persistence may be taken to be the estimated coefficient ζ: the results, together with the estimated standard error, are reported in Table 3. The results confirm that the main difference among the models is between GARCH, TGARCH, and FIGARCH on the one hand and the ANST–GARCH on the other, and that the possibility of switching between regimes entails a drastic reduction in the estimated persistence.

We can therefore expect that the latter model may capture some episodes that the other models do not, and that the former models may exhibit some similarities in their forecasting behavior.

Model	Estimate	Std. Error
GARCH	0.9805	0.0031
TGARCH	0.9777	0.0033
FIGARCH	0.9592	0.0044
ANST-GARCH	0.6164	0.0122

Table 3: Coefficient $\hat{\zeta}$ in the regression $h_t^i = c + \zeta h_{t-1}^i + u_t$

IV The VXN Index

A common problem encountered in estimating volatility models is the way to evaluate relative merits of each for practical purposes. Among the various measures of volatility available to practitioners, one must mention the implied volatilities derived from the prices on the options written on the index itself. As there exists several contracts with different maturities and different strike prices, one interesting aspect which justified the analysis of NDX (NASDAQ–100) is that the implied volatilities of the options written on the index are summarized in a specific index called VXN, introduced in 1995. The VXN is an index of volatility supplied by the Chicago Board of Options Exchange (CBOE).[3] It is computed by taking a weighted average of the implied volatilities of eight call and put NDX options, whose average residual time–to–maturity is 30 days. The VXN consequently can be seen as a market measure of the volatility of the NDX on a 30 day horizon, expressed as the (average) implied volatility on a hypothetical at–the–money option whose time–to–maturity is 30 days. The behavior of the VXN Index from January 3, 1995 to January 18, 2002 (a total of 1839 observations) is shown in Figure 5 (left–hand scale), together with the behavior of the underlying NASDAQ–100 Index (NDX, right–hand scale). In order to facilitate the appreciation of the correspondence between the behavior of the underlying index and the volatility index, we have added some shadings in correspondence with some specific periods of market turbulence. The days were chosen as approximate starting and ending dates on the basis of a sharp change in the VXN Index: Oct. 1, to Oct. 16 1998 – Asian crisis; March 3 to June, 2, 2000 – inception of the "techno–bubble burst", followed by an acceleration of the down turn from Nov. 8, 2000 to May 4, 2001; and, finally, the Sep. 11 crisis following the terrorist attacks – Sep. 17 when markets reopened to Oct. 23, 2001.

The apparent increasing trend in the VXN Index is a clear reflection of the uncertainty that surrounded first the sharp increase in the NDX in the second half of 1999, and, then, the consequences of the sharp decrease of the index in search of new equilibrium values. As the phase, at the time of writing this paper, is not over yet, one cannot speculate (no pun intended!) as of when this increasing trend in the market measure of volatility will end.

[3] See, for further details, the URL *http://www.cboe.com/MktData/vix.asp#vxn*.

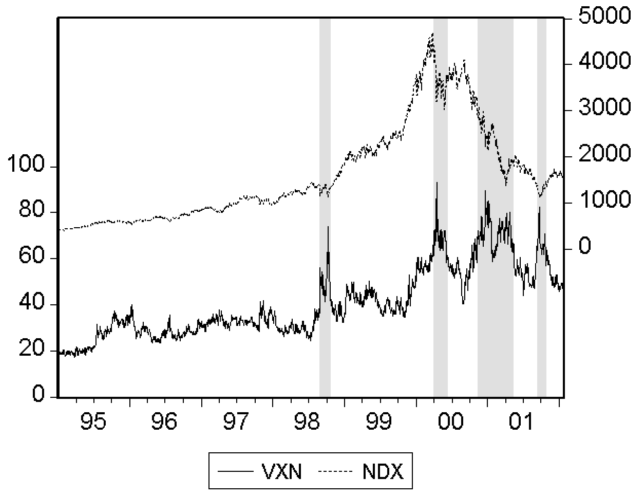

Figure 5: VXN Index and the underlying NDX Index from January 3, 1995 to January 18, 2002

Some descriptive statistics of the VXN Index are reported in Table 4 showing a number of interesting features: while normality is not to be expected given the nature of the series, the presence of autocorrelation and of GARCH effects signals some pattern to be taken into account when modelling the behavior of the series. The research question we are interested in relates to the capability of the GARCH–family forecasts to capture some of the features of this series, a point we will turn to in the next Section.

Mean	40.749	
Median	35.555	
Maximum	93.170	
Minimum	17.360	
Std. Dev.	15.304	
Skewness	0.778	
Kurtosis	2.745	
ADF(4)	-2.48	
Critical Value (5%)	-2.86	
Test	Statistic	p–value
AR(12)	1806.36	0.000
ARCH(4)	1677.50	0.000
Normality	190.61	0.000

Table 4: Descriptive statistics of the VXN Index

V A Forecasting Comparison

On the basis of what we have learned so far, the information content of GARCH model based forecasts should be useful in predicting the VXN series beyond what the past of the series itself may entail. In order to achieve such a comparison, we will need to derive 30–day ahead volatility forecasts using each of the estimated models. The goal is twofold: first, we want to show that the estimated volatility using GARCH models on the NASDAQ–100 Index have predictive power for the (implied–based) volatility index; second, we want to investigate whether the comparison in a combination of forecasts framework will shed some light on the similarities and differences among models. Following Granger and Ramanathan (1984) we will set up a regression where the left–hand side variable is the observed VXN and the regressors are the averaged dynamic forecasts from the models previously analyzed and an AR(1) correction to take into consideration the autocorrelation effects seen before. If there is no information content in the forecasts, the null hypothesis of no impact (zero coefficients) for the regressors on the VXN should be accepted. The rejection of the null, on the contrary, will provide some information about the the suitability of GARCH models to reflect the volatility evaluations expressed by the market.

The predicted series, obtained by averaging the dynamic forecasts 22 steps ahead (corresponding to 30 days) of the conditional variance according to each model presented in Table 2, annualizing the results and taking square roots to get predicted volatilities, are shown in Figure 6.

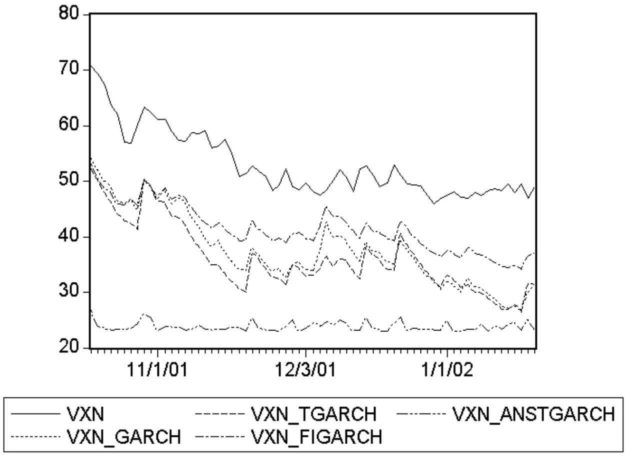

Figure 6: Forecasts of the VXN Index obtained by the volatility models presented in Section 2; October 18, 2001 to January 18, 2002

Graphically, irrespective of the scale differences (which will translate into the regression coefficients being different than one and/or in a positive constant) we can remark that the volatility models follow the behavior of the VXN Index accurately; The striking aspect is, as previously noted in section 3, that the predictions obtained by the GARCH, TGARCH and FIGARCH models are in line with one another, and follow the VXN pattern fairly closely, while the forecasts derived by the ANST–GARCH model seem to have less explanatory power.

The forecast combination regression is estimated over the period Oct. 18, 2001–Jan. 18, 2002. The results are presented in Table 5 with estimated coefficients and (Newey–West autocorrelation robust) standard errors in parentheses. We have regressed the VXN Index on each forecast first, in order to gain some insights on the characteristics of each model. Among the four models taken separately, both the TGARCH and the FIGARCH have regression coefficients which are not significantly different than one (at all customary significance levels), the GARCH coefficient is significantly different only at 1%, while the standard error of the coefficient for the ANST–GARCH is so wide that the confidence interval around the estimated coefficient is very large.

The regression with all four regressors (last column) presents the interesting feature that no coefficient is individually significant while the regression as a whole is highly significant as indicated by the F-test. There are still some problems in the residuals both for autocorrelation and ARCH effects, but the symptoms are clearly pointing to the presence of multicollinearity among the regressors.

Constant	23.783	22.846	14.224	4.656	39.527
	(3.452)	(3.513)	(24.821)	(5.978)	(17.976)
GARCH	0.773	–	–	–	1.654
	(0.092)				(0.591)
TGARCH	–	0.833	–	–	0.079
		(0.095)			(0.446)
ANST–GARCH	–	–	1.628	–	0.414
			(1.054)		(0.413)
FIGARCH	–	–	–	1.165	-1.489
				(0.147)	(0.880)
R–squared	0.775	0.744	0.053	0.747	0.787
Durbin–Watson	0.637	0.674	0.157	0.631	0.641
F–stat	224.03	189.64	3.703	192.79	57.503
AR(4)	29.682	30.323	52.693	30.947	29.298
ARCH(2)	11.489	16.699	46.388	16.450	4.755

Table 5: Regression of the VXN Index on the volatility forecasts from GARCH-type models (robust standard errors in parentheses)

The correlation matrix among the forecasts (cf. Table 6) shows that the linear correlation between pairs of GARCH, TGARCH and FIGARCH is very high, thus confirming the diagnostics on the parameter significance problems pointed out before. One of the possible remedies is the adoption of a principal component approach in order to recombine the forecasts together and point out what orthogonal explanations of variation in the VXN can be derived from the data. By adopting a principal component extraction, as done in Table 7, we see that the first two components are capable of explaining in excess of 98% of the total variance. The eigenvectors associated to the eigenvalues (not shown here) suggest (not surprisingly) that the first component is related to the three models GARCH, TGARCH and FIGARCH, while the second principal component is linked to the outcome obtained by the ANST–GARCH model.

With these results in hand, we are equipped to perform a regression in which the regressors are the principal components themselves and a constant. The results are shown in Table 8 where we have inserted the third component as well in order to underline its lack of significance. However, a disturbing feature of the regression diagnostics is that the residuals are affected by autocorrelation and ARCH effects. A solution is one in which we insert the lagged VXN in the regression, in view of the high persistence of the index highlighted previously. In the third column of the Table 8, therefore, we report the estimated coefficients of the resulting equation. The coefficient on the principal component related to the ANST-GARCH is not significant, while the residual autocorrelation and ARCH effects have disappeared. Finally, therefore, we prune the corresponding variable from the regression and display the results in the fourth column. As one would expect, there is no consequence on the residual properties from doing so.

The fitted values and the residuals of the second and fourth regression are displayed, respectively, in Figures 7 and 8: the greater suitability of the latter is apparent.

	GARCH	TGARCH	ANST–GARCH
TGARCH	0.9652		
ANST–GARCH	0.2623	0.3253	
FIGARCH	0.9939	0.9564	0.3149

Table 6: Correlation matrix between forecasts

Component	Eigenvalue	% of variance
PC1	3.07472	76.868 %
PC2	0.87196	21.799 %
PC3	0.04980	1.245 %
PC4	0.00352	0.088 %

Table 7: Principal component decomposition of the forecasts of the VXN Index

VI Concluding Remarks

In this paper we have addressed some issues arising in using GARCH–type models in forecasting conditional variance. The suitability of these models in capturing various features of observed financial time series (particularly volatility clustering and fat–tailed distributions) is well documented and is reproduced here in specific reference to the NASDAQ–100 Index. Pushing the argument one step forward, we have investigated whether these models provide *ex–ante* forecasts which are related to the market valuations of volatility as reflected in the implied volatilities derivable from option prices. The interest in the NASDAQ–100 Index is that we have available a derived volatility index, the VXN, related to it, calculated as an average of the implied volatilities of put and call options with residual maturities close to 30 days. By calculating 30–day ahead

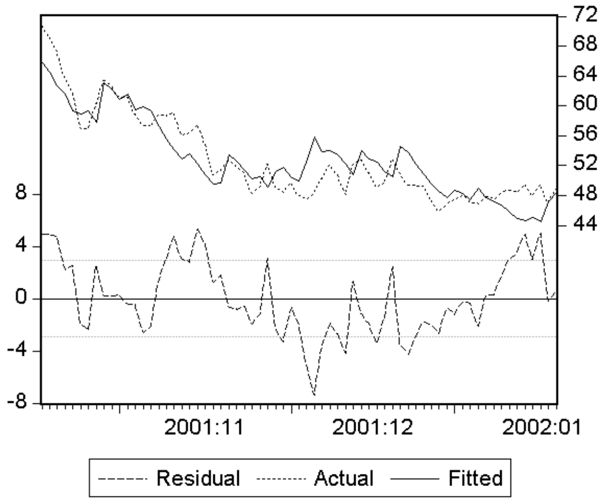

Figure 7: Actual, fitted, residual graph for the regression of VXN on a constant and the first two principal components (coefficients in the second column of Table 8)

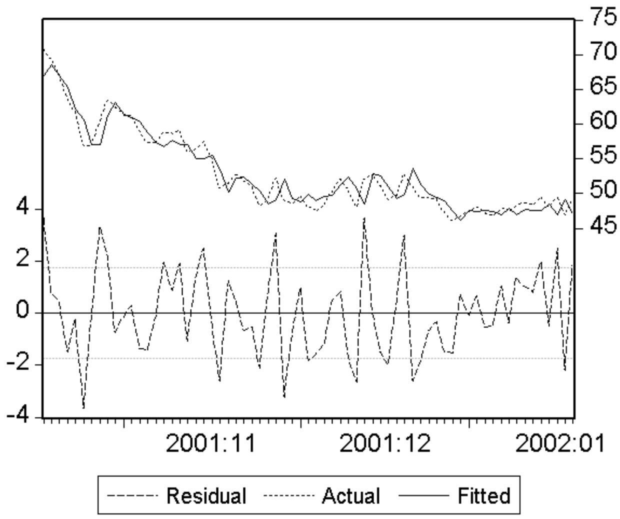

Figure 8: Actual, fitted, residual graph for the regression of VXN on a constant, lagged VXN and the first principal components (coefficients in the fourth column of Table 8)

Constant	52.964	52.964	12.304	13.671
	(0.358)	(0.355)	(2.881)	(2.974)
VXN_{t-1}			0.763	0.738
			(0.054)	(0.055)
PC1	2.933	2.933	0.673	0.749
	(0.226)	(0.225)	(0.195)	(0.240)
PC2	0.969	0.969	-0.211	
	(0.378)	(0.375)	(0.264)	
PC3	-0.336			
	(2.123)			
R-squared	0.7718	0.7716	0.9203	0.919
Durbin-Watson	0.6589	0.6567		
F-stat	**71.029**	**108.13**	**242.68**	**365.33**
AR(4)	**30.298**	**29.984**	9.229	5.540
ARCH(2)	**10.388**	**10.667**	0.503	0.424

Note that the coefficients in the first and second column are the same as the regressors are orthogonal.
In parentheses we report Newey West robust standard errors.
Test AR is Breusch–Godfrey $TR^2 \sim \chi^2_4$ from the auxiliary regression;
Test ARCH is Engle test $TR^2 \sim \chi^2_2$ from the auxiliary regression;
In boldface values of the test statistics significant at 5% level.

Table 8: Regression of the VXN Index when principal components of the original forecasts are used.

forecasts (22 trading days) of the volatility on the NASDAQ–100 Index (thus ignoring the information available on the VXN) we have mimicked the behavior of the volatility. In a forecasting combination exercise this strategy pays off by showing that, beyond what is given by the past of the volatility index, the GARCH–based volatility forecasts provide useful insights on where the index is aiming at.

The next question we have investigated is about the relative merits of the various models. In spite of the sample chosen for the *ex–ante* forecasting exercise which is limited to a few months, the results show clearly that both in estimation and forecasting, the GARCH, TGARCH and FIGARCH models behave similarly to one another. The estimation of the ANST–GARCH produced markedly different results, as expected, given that the presence of regimes hypothesized by the model reduces the persistence of the estimated volatility. These similarities and differences are reproduced in the combination of forecasts exercise: high correlation among the three models translated into collinearity problems and therefore suggested the adoption of a principal component strategy to extract orthogonal information. In this respect the first principal component is strongly connected with the three models and the second with the ANST-GARCH model. However, in order to avoid autocorrelation and residual ARCH effects, upon insertion of the lagged value of the VXN Index, the significance of the second component disappeared.

We need to stress that the sample chosen for the forecasting combination comparison is a fairly peculiar one, since it immediately follows the aftermath of the September

11, 2001 terrorist attacks and the overall uncertainty surrounding the state of the US economy at the turn of the year. The performance of the ANST–GARCH is somewhat disappointing, since it seems to produce less useful forecasts: in its favor, though, one should note that it is possible that the latest period is one in which the features of the regimes as estimated in–sample have changed thus forcing the ANST–GARCH model to operate in an unexplored territory.

Overall, the results can be deemed very satisfactory, and should be interpreted as a positive feature of the standard GARCH models being able to appropriately capture the volatility dynamics even in times of turbulence.

References

Andersen, T.G., Bollerslev, T. (1997): Heterogeneous information arrivals and returns volatility dynamics: uncovering the long run in high frequency returns.*Journal of Finance*, *52*, 975–1005.

Anderson, H.M., Nam, K., Vahid, F. (1999): Asymmetric Nonlinear Smooth Transition GARCH Models, in P. Rothman (ed.), *Nonlinear Time Series Analysis of Economic and Financial Data,* Boston: Kluwer, 191–207.

Baillie, R.T., Bollerslev, T., Mikkelsen, H.O. (1996): Fractionally Integrated Generalized AutoRegressive Conditional Heteroskedasticity. *Journal of Econometrics*, *74*, 3–30.

Bollerslev, T. (1986): Generalized AutoRegressive Conditional Heteroskedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollerslev, T., Chou, R.Y., Kroner, K.F. (1992): ARCH Modeling in Finance. *Journal of Econometrics*, *52*, 5–59.

Bollerslev, T., Engle, R.F., Nelson, D. (1994): *ARCH Models,* in McFadden, D.L. and Engle, R.F. (editors), *Handbook of Econometrics,* Vol. 4 Chapter 49, 2959–3038.

Caporin, M. (2001): Estimation and Identification of FIGARCH. *Unpublished working paper, Università "Ca' Foscari" di Venezia*.

Davidson, J.E. (2001): Moment and Memory Properties of Linear Conditional Heteroscedasticity Models. *Unpublished working paper, Cardiff University*.

Ding, Z.-X., Granger, C.V.J., Engle, R.F. (1993): A Long Memory Property of Stock Market Returns and a new Approach. *Journal of Empirical Finance*, *1*, 83–106.

Engle, R.F. (1982): AutoRegressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, *50*, 987–1008.

Engle, R.F., Gallo, G.M. (2001): A Model for Intra-Daily Volatility with Multiple Indicators. *Unpublished working paper, UCSD, San Diego*.

Engle, R.F., Patton, A.J. (2001): What Good is a Volatility Model? *Quantitative Finance*, *1*, 237–245.

Glosten, L.R., Jagannathan, R., Runkle, D.E. (1993): On the Relation between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks. *Journal of Finance*, *48*, 1779-1801.

González–Rivera, G. (1998): Smooth Transition GARCH Models. *Studies in Nonlinear Dynamics and Econometrics*, *3*, 61–78.

Granger, C.W.J., Ramanathan, R. (1984): Improved Methods of Combining Forecasting, *Journal of Forecasting*, *3*, 1984, 197-204

Gray, S. F. (1996): Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. *Journal of Financial Economics*, *42*, 27-62.

Hagerud, G.E. (1997): *A New Non–Linear GARCH Model*, PhD thesis, Stockholm School of Economics.

Hamilton, J.D., Susmel, R. (1994): Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics*, *64*, 307–333.

Hansen, B.E. (1996): Inference when a Nuisance Parameter is not Identified Under the Null Hypothesis. *Econometrica*, *64*, 413–430.

Lombardi, M.J., Gallo, G.M. (2001): Analytic Hessian Matrices and the Computation of FIGARCH Estimates. *Working Paper 2002/03, Dipartimento di Statistica "G. Parenti", Università degli studi di Firenze*.

Teräsvirta, T. (1994): Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. *Journal of The American Statistical Association*, *89*, 208–218.

Zakoian, J.–M. (1994): Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, *18*, 931-955.